Computational Science Working Group

Adam Lyon & Jim Kowalkowski

All Scientist Retreat

26 April 2018

Micro-workshop https://indico.fnal.gov/event/16923/

Friday, 20 April 2018

13:30 - 13:40 Introduction

13:30 **Why are we here?** *10'*
Speaker: Dr. Adam Lyon (Fermilab)
Material: **Slides**

13:40 - 14:55 Stepping stones towards DUNE/HL-LHC (near and farther future)

13:40 **CWP and S2I2** *5'*
Speaker: Dr. Rob Kutschke (Fermilab)
Material: **Slides**

13:45 **HEPCloud Direction** *5'*
Speaker: Dr. Burt Holzman (FNAL)
Material: **Slides**

13:50 **Machine Learning Directions** *5'*
Speaker: Dr. Brian Nord (Fermilab)
Material: **Slides**

13:55 **ROOT i/o directions** *5'*
Speaker: Mr. Philippe Canal (FERMILAB)
Material: **Slides**

14:00 **Reconstruction directions** *5'*
Speaker: Giuseppe Cerati (Fermilab)
Material: **Slides**

14:05 **Simulation Directions** *5'*
Speaker: Soon Yung Jun (Fermilab)
Material: **Slides**

14:10 **CMS Directions** *15'*
Speakers: Dr. Oliver Gutsche (Fermi National Accelerator Laboratory), Dr. Christopher Jones (Fermilab), Nhan Tran (FNAL)
Material: **1. CMS R&D Overview (PDF)** **Additional Material for CMS R&D Overview** **Slides**

14:25 **Framework Directions** *20'*
Speakers: Dr. Kyle Knoepfel (Fermilab), Dr. Marc Paterno (Fermilab), Saba Sehrish (Fermilab), Kowalkowski (Fermilab)
Material: **Slides**

14:45 **Heterogenous Computing and Coherent Interconnects** *5'*
Speaker: Dr. Michael Wang (Fermilab)
Material: **Slides**

14:50 **Cosmic Software Directions** *5'*
Speaker: Dr. James Annis (Fermilab)
Material: **Slides**

Problems to Confront (~2026+)

14:55 **Post Moore Introduction** *10'*
Speaker: Mr. Jim Kowalkowski (Fermilab)
Material: **Slides**

15:05 **Computing Security** *5'*
Speaker: Dr. Mine Altunay (FNAL)

15:10 **DAQ Directions** *5'*
Speaker: Dr. Gustavo Cancelo (fermilab)
Material: **Slides**

15:15 **TDAQ Workshop Summary** *5'*
Speaker: Mr. Alan Prosser (Fermilab)
Material: **Slides**

15:20 **Energy Frontier Directions** *5'*
Speaker: Dmitri Denisov (Fermilab)
Material: **Slides**

15:25 **More for CMS** *5'*
Speaker: Lindsey Gray (Fermilab)
Material: **Slides**

Watch the movie!

Charge

- What are the interests of the Fermilab scientists for the decade or so following 2026?
- How do we give our input to both the US community planning and the European Strategy Group?
- What is the post-retreat plan for working with US, European, and other partners to give our input?

Addressing 2nd and 3rd charge items…

HEP-CCE: Promoting Computational Excellence

HEP-CCE Coordinators: Salman Habib (Argonne), Kerstin Kleese Van Dam (Brookhaven), Rob Roser (Fermilab), and Peter Nugent (Lawrence Berkeley)

The HEP-CCE is a cross-cutting initiative to promote excellence in high performance computing (HPC) including data-intensive applications, scientific simulations, and data movement and storage. Enhancing connections with DOE's Advanced Scientific Computing Research (ASCR) program is an important part of the Center's activities.

HSF
HEP Software Foundation

The HEP Software Foundation facilitates cooperation and common efforts in High Energy Physics software and computing internationally.

Structures are in place to identify cross-cutting R&D opportunities and advise funding agencies — Fermilab has important involvement

Community White Paper (CWP)

International effort to determine R&D Roadmap for HL-LHC (and DUNE)

Stewarded by HSF — Fermilab input into nearly all reports

Community White Paper Reports

The roadmap summarised reports from fourteen working groups who studied the challenges in their sub-domains. All of the reports produced during the Community White Paper process are listed below. Working groups are in the process of finalising and uploading their work to arXiv.

Paper	*Report Number*	*Link*
CWP Roadmap	HSF-CWP-2017-01	arXiv
Careers & Training	HSF-CWP-2017-02	ShareLaTeX
Conditions Data	HSF-CWP-2017-03	Google Doc
Data Organisation, Management and Access	HSF-CWP-2017-04	Overleaf
Data Analysis and Interpretation	HSF-CWP-2017-05	arXiv
Data and Software Preservation	HSF-CWP-2017-06	Google Doc
Detector Simulation	HSF-CWP-2017-07	arXiv
Event/Data Processing Frameworks	HSF-CWP-2017-08	Google Doc
Facilities and Distributed Computing	HSF-CWP-2017-09	Google Doc
Machine Learning	HSF-CWP-2017-10	ShareLaTeX
Physics Generators	HSF-CWP-2017-11	Overleaf
Security	HSF-CWP-2017-12	See section 3.13 of roadmap
Software Development, Deployment and Validation	HSF-CWP-2017-13	arXiv
Software Trigger and Event Reconstruction	HSF-CWP-2017-14	arXiv - Executive Summary; arXiv - full document
Visualisation	HSF-CWP-2017-15	Google Doc

arXiv.org > physics > arXiv:1712.06982

Search or Artic

(Help | Advanced s

Physics > Computational Physics

A Roadmap for HEP Software and Computing R&D for the 2020s

HEP Software Foundation: Johannes Albrecht, Antonio Augusto Alves Jr, Guilherme Amadio, Nguyen Anh-Ky, Laurent Aphecetche, John Apostolakis, Makoto Asai, Luca Atzori, Marian Babik, Giuseppe Bagliesi, Marilena Bandieramonte, Sunanda Banerjee, Martin Barisits, Lothar A. T. Bauerdick, Stefano Belforte, Douglas Benjamin, Catrin Bernius, Wahid Bhimji, Riccardo Maria Bianchi, Ian Bird, Catherine Biscarat, Jakob Blomer, Kenneth Bloom, Tommaso Boccali, Brian Bockelman, Tomasz Bold, Daniele Bonacorsi, Antonio Boveia, Concezio Bozzi, Marko Bracko, David Britton, Andy Buckley, Predrag Buncic, Paolo Calafiura, Simone Campana, Philippe Canal, Luca Canali, Gianpaolo Carlino, Nuno Castro, Marco Cattaneo, Gianluca Cerminara, Javier Cervantes Villanueva, Philip Chang, John Chapman, Gang Chen, Taylor Childers, et al. (250 additional authors not shown)

(Submitted on 18 Dec 2017 (v1), last revised 11 Feb 2018 (this version, v3))

Particle physics has an ambitious and broad experimental programme for the coming decades. This programme requires large investments in detector hardware, either to build new facilities and experiments, or to upgrade existing ones. Similarly, it requires commensurate investment in the R&D of software to acquire, manage, process, and analyse the shear amounts of data to be recorded. In planning for the HL-LHC in particular, it is critical that all of the collaborating stakeholders agree on the software goals and priorities, and that the efforts complement each other. In this spirit, this white paper describes the R&D activities required to prepare for this software upgrade.

Subjects: **Computational Physics (physics.comp-ph)**; High Energy Physics - Experiment (hep-ex)
Report number: HSF-CWP-2017-01
Cite as: arXiv:1712.06982 **[physics.comp-ph]**
(or arXiv:1712.06982v3 **[physics.comp-ph]** for this version)

Submission history

From: Graeme Stewart [view email]
[v1] Mon, 18 Dec 2017 17:55:32 GMT (310kb,D)
[v2] Wed, 20 Dec 2017 08:33:04 GMT (310kb,D)
[v3] Sun, 11 Feb 2018 19:38:37 GMT (283kb,D)

Fermilab

S2I2 —> IRIS

Community White Paper reports inform NSF on establishing a software institute

Scientific Software Innovation Institute (S2I2) —> Institute for Research in Innovative Software (IRIS)

NSF funded; Lead by Peter Elmer, Mark Neubauer, Mike Sokoloff

Delayed start due to budget uncertainty
Will focus on HL-LHC Software R&D
Coincidental overlap with neutrino/muon needs may be exploited

DOE Funds Computing R&D…

OHEP with COMPHEP and CCE

- Detector Simulations (Geant)
- Accelerator Simulations
- Software Frameworks including new architectures
- Big Data & Machine Learning
- Running on HPC (Supercomputers)
- Lattice QCD (joint with NP)
  Problem well suited to early adoption of HPC technology
- CMS Computing & Software R&D

ASCR Office (Advanced Scientific Computing Research)

- Operates HPC centers (ALCF, OLCF, NERSC)
- HPC R&D

ASCR funding to HEP…

- SciDAC (Scientific Discovery though Advanced Computing)
  - $17.5M awarded to FNAL: two 5 year projects and one 3 year project
  - Accelerator Modeling (5yr)
    Reconstruction on advanced architectures (3yr)
  - HEP Data Analytics on HPC
    LHC/Neutrino Science, Optimization, Storage and Data Modeling, Workflow (5yr)
- Exascale Computing Project for Lattice QCD (joint with BNL, JLab)

LDRD: Off-the-shelf DAQ; Databases for Big Data; HEP with Micron Automata; Preparing HEP for Exascale, QC, ML

HEP Data Analytics on HPC SciDAC (JBK)

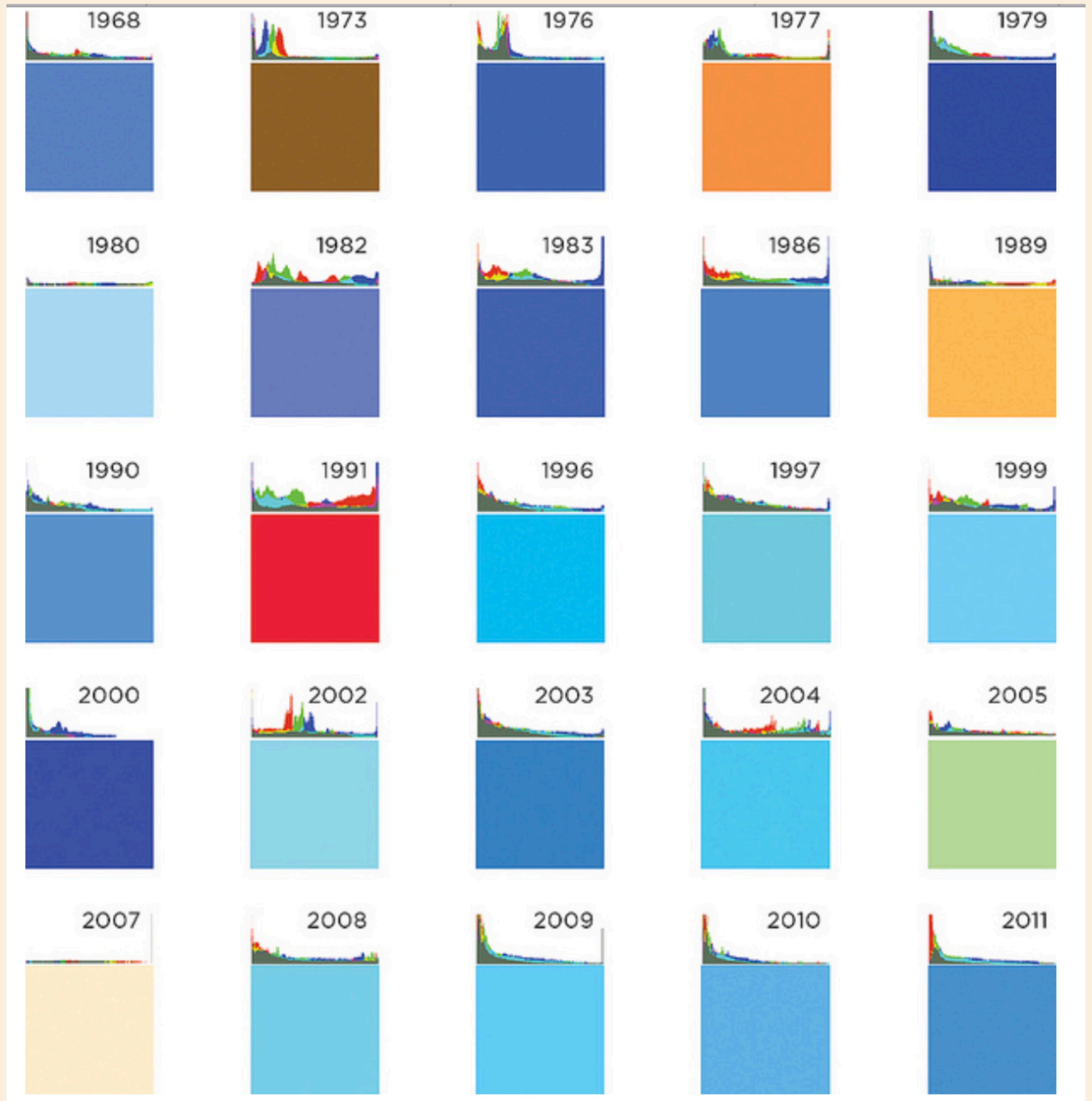
Crude timeline

CMS/ATLAS Generator Tuning	Current generation analysis – new tools	Next generation techniques with automation	
NOvA/DUNE Oscillation/Cross-section analysis	Use existing code and tools	Oscillation analysis Next Generation	Cross-section analysis

- Focus on whole-dataset analysis, utilizing whole machine
  - Introduce optimization and workflow tools from HPC community
  - Incorporate simulation, reconstruction, and selection layers into analysis procedures
- Generator analysis and tuning will lead in development and utilization of new tools and techniques
- Experiment analysis will start with existing codes and application to establish baselines, next generation systems will adapt and utilize techniques from generator groups

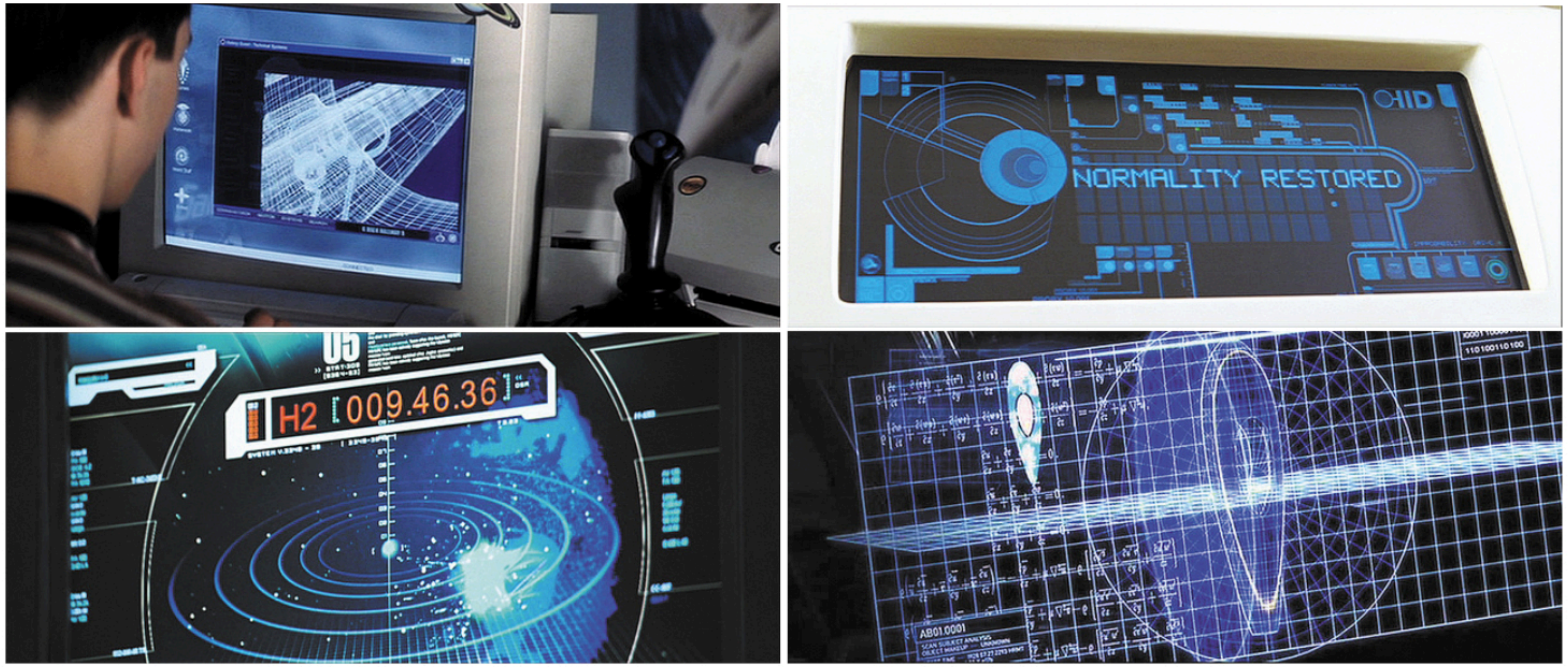
What will Computing Look Like > 2026?

SciFi says your screen will be blue (unless you are a terminator )



1991 was Terminator 2. Chart courtesy of Nathan Shedroff and Chris Noessel

Make It So: Interaction Design Lessons from Science Fiction



Clockwise from Left: 1. Galaxy Quest (1999) Dreamworks; 2. The Hitchhiker's Guide to the Galaxy (2005) Touchstone; 3. Supernova (2000) MGM; 4. Fantastic 4 (2005) 20th Century FOX

https://99percentinvisible.org/episode/future-screens-are-mostly-blue/

Fermilab

What will Computing Look like > 2026?

We know shorter term, but not long term …
won't try to guess

Instead, think about what we'll be doing in 2026+

How would computing support that science? What are the computing trends?

What R&D would be necessary and make a roadmap.

Three "triggers" for Computational R&D…

1) Receive requirements from experiments based on upcoming needs
2) Forward thinking to keep up with the evolving computing landscape
3) Useful technologies that scientists adopt and needs support

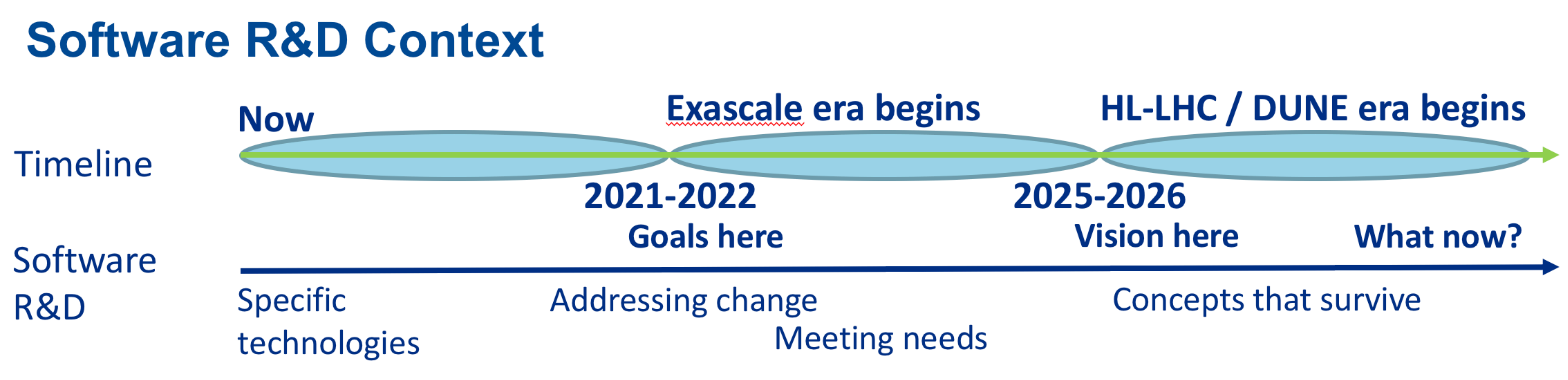
Three areas for R&D

A) Computational Software

B) Operating Computing Systems

C) Data Acquisition

Timeline



JBK

Where is Computing Going?

Moore's Law:

# of transistors doubles every two years

Dennard Scaling:

Power/transistor decreases

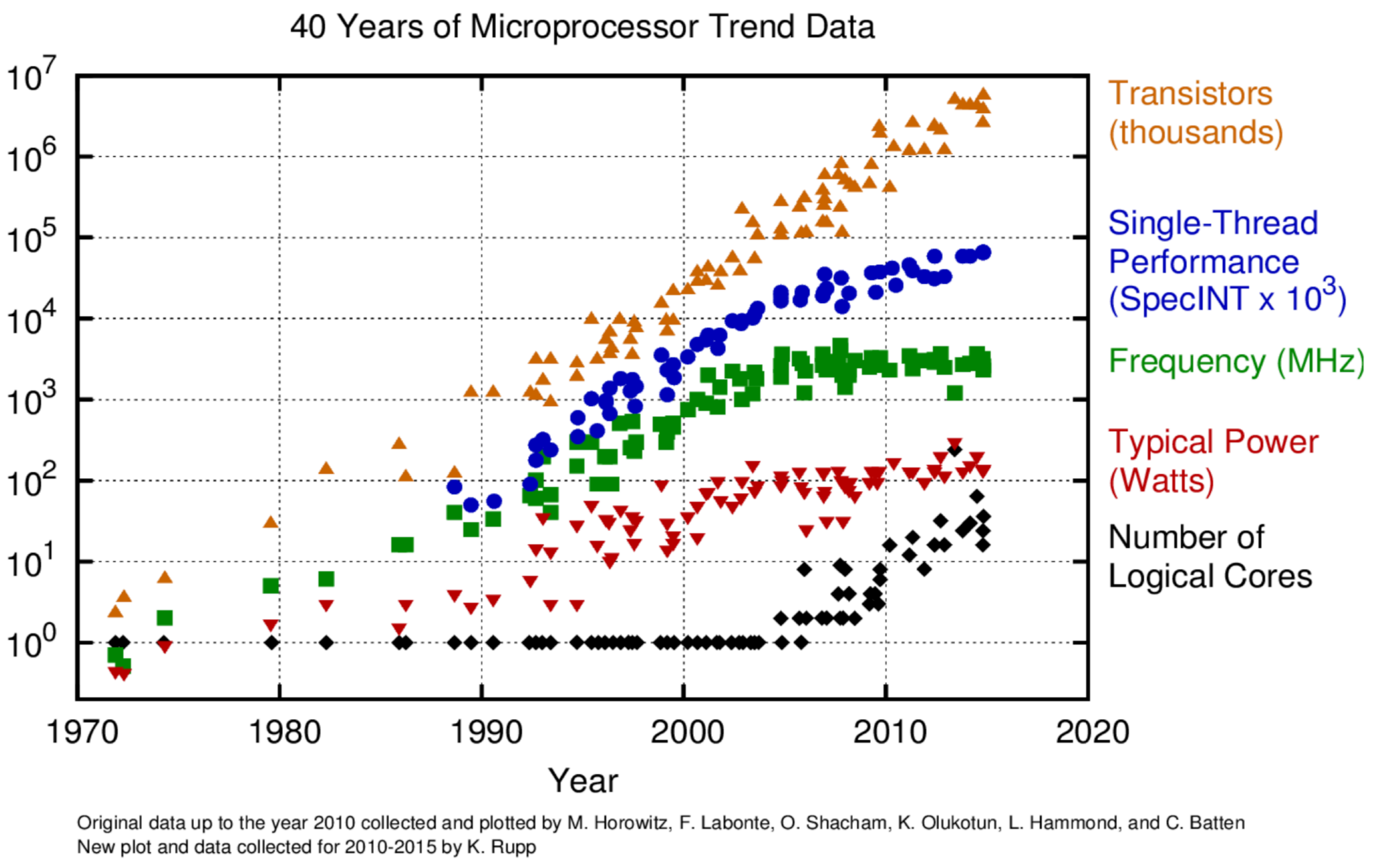
so clock speeds can increase without increasing total power consumed

Clock speeds have been constant for 10 years

Can't make cores faster, so give you more of them

- **Multiprocessors**
- **Multithreading**

I've mentioned R&D already



Fermilab

Exascale Computing

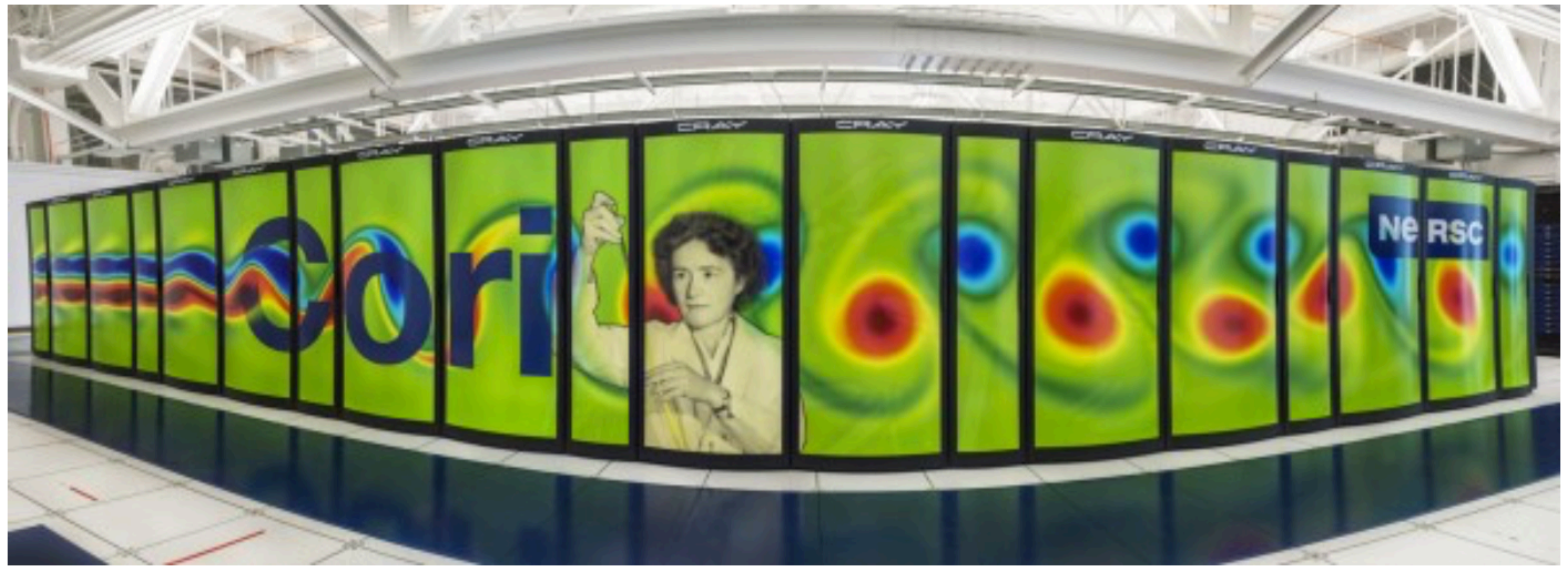
Massively parallel Supercomputers (NSCI/ECP)
[Major challenge is energy efficiency]

CORI (NERSC): 153K Haswell Threads
2.6M KNL Threads

Summit (ORNL): 27K GPUs; 9.2K POWER9
Important for ML training

Aurora (ANL): Was to be next generation KNL
- Now likely an "extreme heterogeneity" machine
- Specialized hardware for Big Data, ML, HPC
- Details yet to be revealed - targeted for ~2021

Much R&D now and short term future to learn how HEP can effectively use these resources (vectorization and multithreading)



Post-Moore Computing

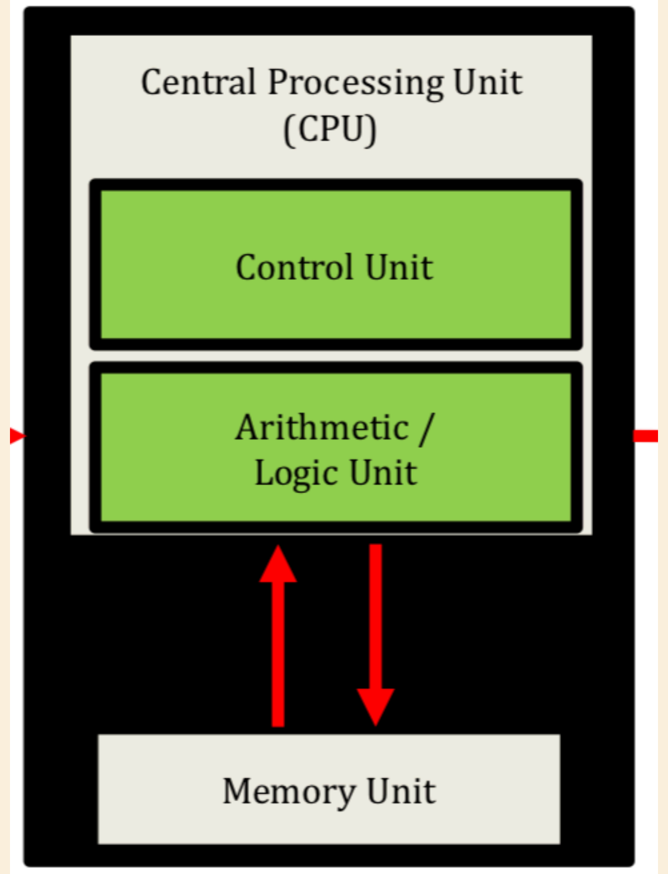
R&D Necessary for HEP to adopt

Reach the limit of # of transistors on a chip (probably around 2020)

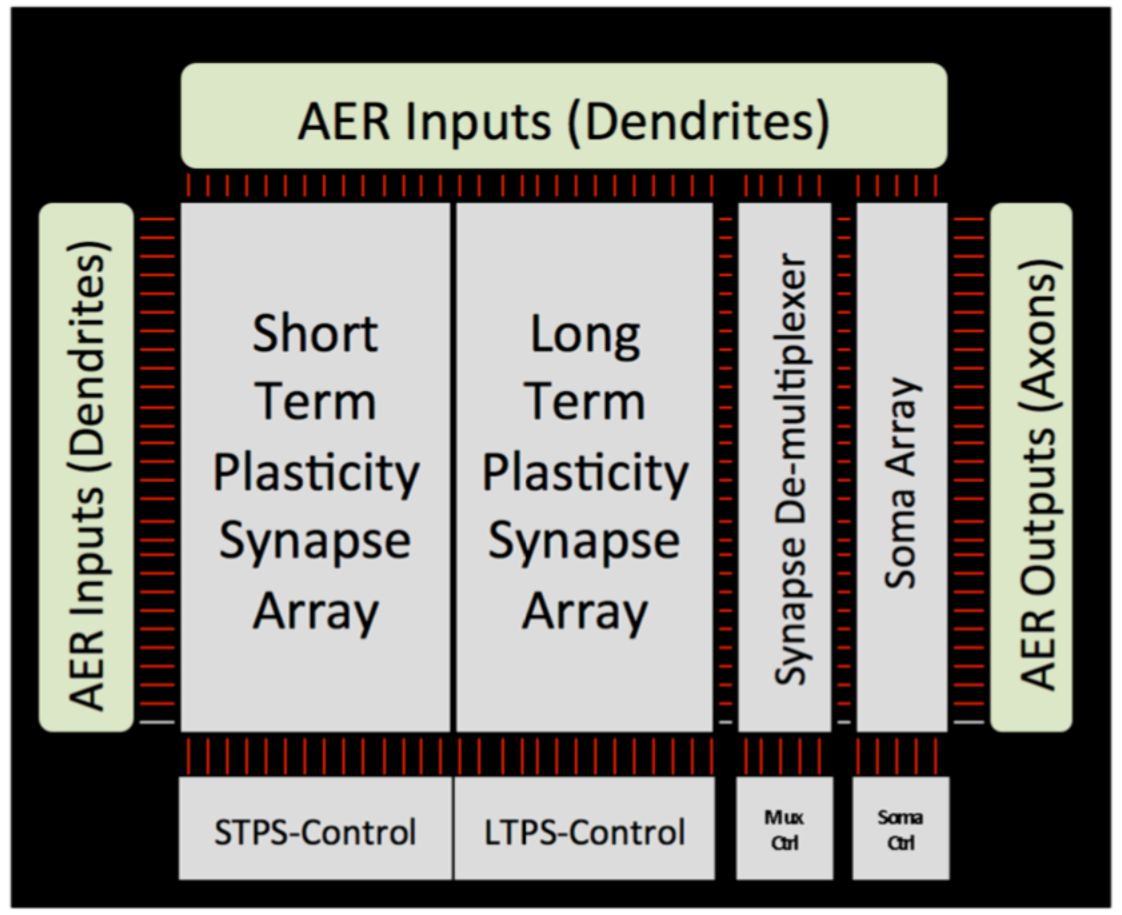
New and different computing emerges — ASCR is driving

- What does it mean? (not mutually exclusive categories)
  - Energy efficient computing
  - Exotic technology
  - Extreme heterogeneous computing (Machines with CPUs, GPUs, TPUs, …)
  - Processing in close proximity to peripheral systems
  - FPGAs everywhere; massive memory replacing massive storage
- Already see evidence of this depending on definition you like best
- Well-known contenders in the exotic technology realm
  - Quantum computers (the latest craze, includes D-Wave)
  - Neuromorphic Computing (C. Shuman gave a few talks here on the subject)
  - Micron’s automata processors (Practically dead)
  - **Shared property:** *Very much unconventional programming here*

JBK

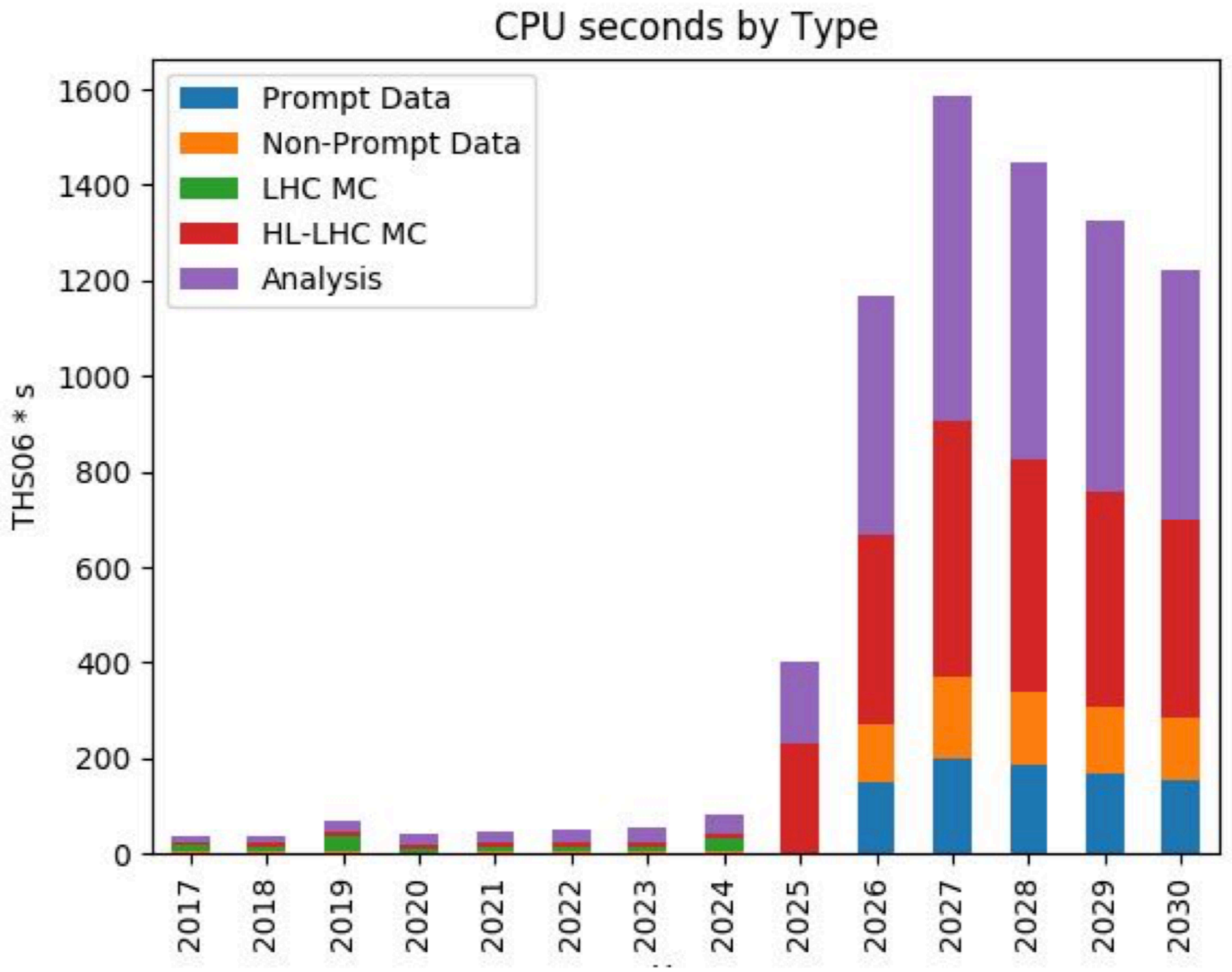


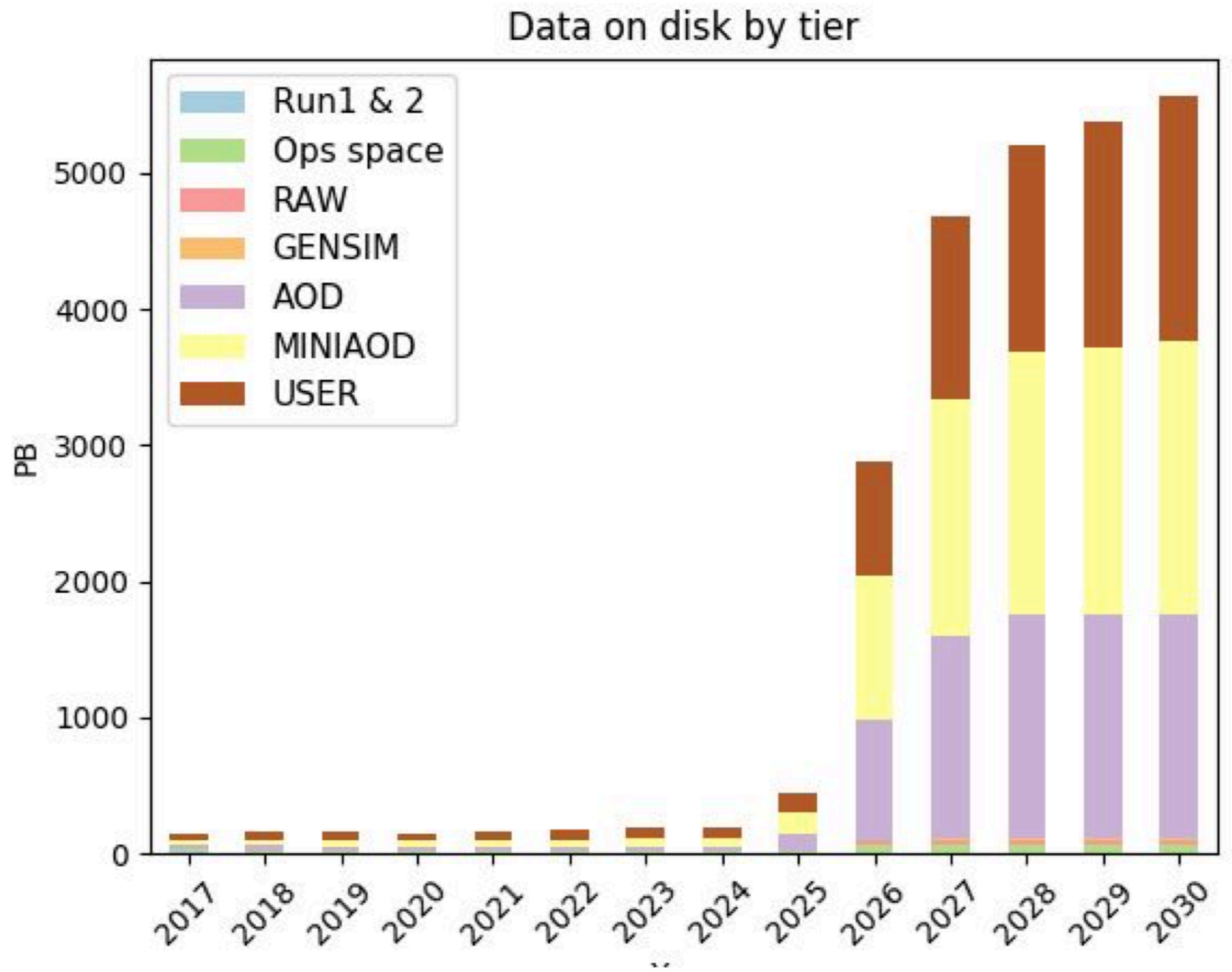
Neuromorphic Architecture



Fermilab

CMS R&D





Naive extrapolation for HL-LHC:

In 2027 we "would" need:

- CPU: 5 Million cores (x20 compared to now)
- Disk: 5 Exabyte (x50 compared to now)

We need:
Disruptive Changes!

Oli

Fermilab

CMS R&D

CMS R&D overview

List is not exhaustive

Computation	**Many-Core** More cores, less memory	**Accelerators** GPUs, FPGAs	**Algorithms** Parallel Kalman, GeantV
Storage	**Store Data** NanoAOD	**Organize Data** Data Lakes	**Analyze Data** Big Data, HPC
Services	**Elasticity** **HEPCloud**	**Connectivity** Software Defined Networking	**Orchestration** Containers

Oli

Other Software R&D

Machine learning for full reconstruction and simulation

Vectorization and parallelization at algorithm level (reco/sim)

Auto optimized code generation for heterogeneous systems

ROOT: Pass through i/o, i/o for parallelization, object stores

Frameworks: reduce dependencies, functional programming, whole-dataset operations, programming/data models NOT tightly coupled to language, tiered memory usage

ML on diverse hardware

FPGAs closely interconnected to CPUs (ML, triggering, reconstruction, analysis)

Worry: What do we do when Quantum Computing breaks all encryption?

Continue our strategy of COMMON TOOLS

Future experiments

Future EF (Higgs factory/100 TeV pp) go far beyond HL-LHC

- The technology needed to step beyond HL-LHC may be a ways off
- R&D for HL-LHC should be a good guide

Future CF (LSST/DESC/CMB-S4)

- Very large data sets; image processing; spatial processing
- Common workflow important

DAQ

Technology Topics for Future Consideration

DOE DAQ Workshop 10/17

1. Common warm RF electronics control/DAQ for future CF experiments
2. Rad-hard, high-speed optical components
3. Wireless communication for data transfers
4. Optimum locations in the data stream for compression & filtering (closer to FE)
5. Coprocessor investigation and development (more from Jim K).
6. Deep learning on FPGAs/High Level Synthesis code methodology

Alan

CCD/MKIDS DAQs — ~0.5M detectors at high rates, warm electronics,
RF controls may be useful for Quantum Computers

How are we moving forward?

Execute the R&D Projects we have now and succeed

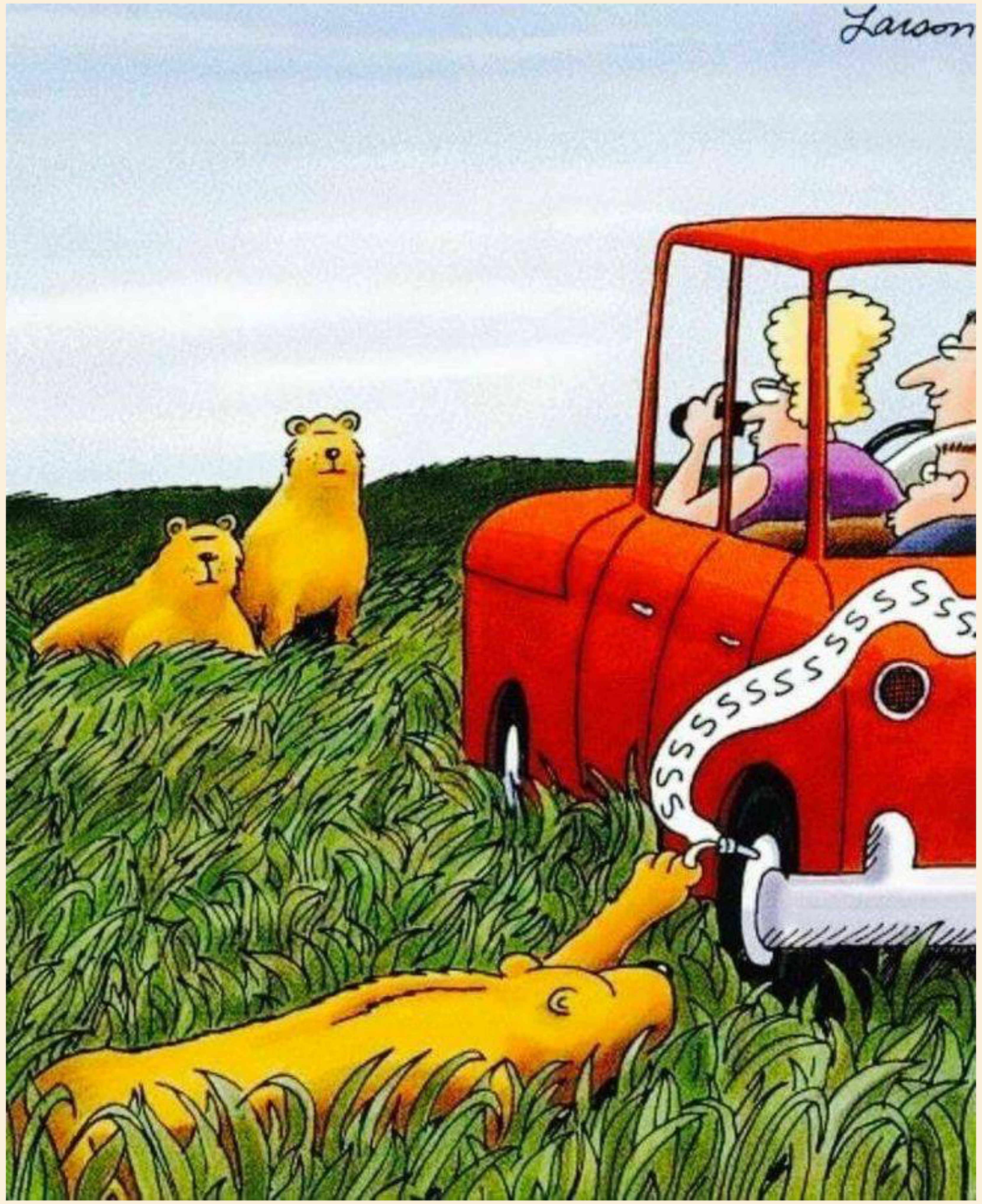
Follow on with new proposals and projects

Continue to engage ASCR (they're driving the paradigm shifts in the US)

Work with our partners and plan the future

- Universities helped by IRIS
- Other labs helped by CCE
- Internationally helped by HSF

Maintain our leadership in HEP Computing R&D

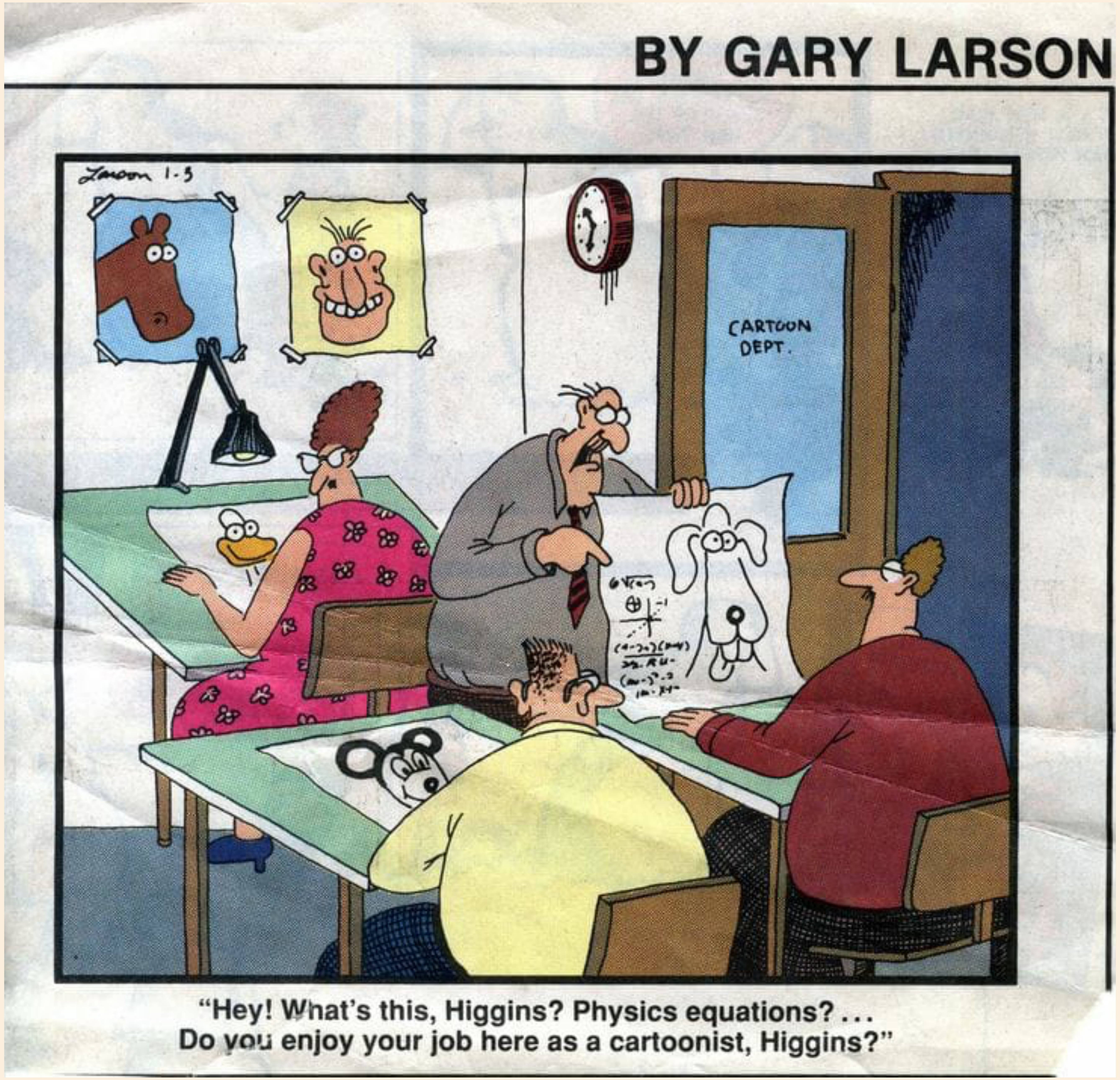


Must be aware of what's happening in the computing neighborhood

Can't let the future get the jump on us

Fermilab

We do Computing R&D to support and enable the Physics



Fermilab